Early-Stage Epidemic Forecasting

Confirmation of the Grey Model Theory to Early Stage Epidemic Forecasting

Olayemi Olabisi
University of Liverpool
Liverpool, England
e-mail: olayemi_olabisi@hotmail.co.uk

Andrea Corradini
Copenhagen School of Design and Technology
Copenhagen, Denmark
e-mail: andc@kea.dk

***Abstract*—Statistical methods and machine learning methods are currently the most popular ways for forecasting. Some of these include autoregressive models, cumulative sum charts, growth models, Support Vector Machines for regression (SVR), Polynomial Neural Networks (PNNs) as well as several others. The inherent limitation associated with these models is that they require large sample size for accurate forecasts. In this paper, we investigate the ability to forecast a disease outbreak when data is limited by employing variations of the Grey Model (GM) forecasting. Lack or limitation of data is rather common in the early stages of a disease outbreak. We present the results of a simulation that shows the model's ability to leverage the exponential growth associated with the rate of spread of diseases in the early stages. A comparative analysis of our quantitative results using the coronavirus dataset of a few countries indicate that both the Gompertz model and the PNN model perform better than the traditional Grey Model GM(1,1) in fitting and forecasting. At the same time, qualitative evidence indicate that the Grey Model is suitable for early-stage epidemics provided methods of enhancement are employed i.e. improved background value calculations and better methods for accumulative generating operation as seen in the Fractional Grey Model FGM(1,1) algorithm.**

Keywords- Forecasting; Grey Forecasting model; Early Stage Epidemic; Coronavirus.

I. INTRODUCTION

On the 31st of December 2019, World Health Organization (WHO)'s country office in the people's republic of China picked up a media statement by the Wuhan municipal health commission from their website on cases of viral pneumonia in Wuhan. On the 1st of January 2020, the WHO activated its incident management support team (IMST) as part of its emergency response framework. On the 9th of January 2020, the WHO reported that the Chinese authorities have confirmed the outbreak of a novel coronavirus. It took eleven days for the confirmation of the outbreak of the virus. Since the official declaration of the novel coronavirus as a global pandemic, the world has witnessed an immense spread of the infection. By early September 2021, the virus has spread to nearly every country causing more than 215 million infection cases and nearly 4.5 million deaths worldwide [1], upending lives and derailing the global economy.

Accurate forecasting is one of the tools that are typically used to plan and mitigate risk after and during the outbreak of a disease. Its importance is something we have come to appreciate in the recent coronavirus outbreak. Unlike normal forecasting, epidemic forecasting exhibits traits which have been described in [2][3] especially when considering the initial period of the outbreak of a disease. This initial period can be identified by a sharp rise in cases where numbers are initially dormant. An example of this would be in cases of influenza, where at the beginning of the season little or no cases are observed, then a sudden surge in number of cases is observed [3].

There are several models, which can be employed for modelling and forecasting the initial stage of an epidemic. Examples of models used in previous studies include autoregressive integrated moving average, holt-winters multiplicative, linear regression, support vector machine for regression, fast decision learner, regression trees, PNNs and PNN+cf [4]. Other proposed models for forecasting disease outbreaks include generalized growth models as described in [5]. The problem inherent with these methods and with other statistical and machine learning techniques is that they require large datasets for higher levels of accuracy [6]. In the coronavirus case, it took eleven days for the virus to be identified as a threat and as such only eleven data points were available.

Small sample set problems [7][8] refer to the cases where the number of samples is less than 50 in respect to engineering applications or less than 30 regarding academic research. Such small number sample size cannot fully reveal the behavior of a system or population [9]. There are three main approaches for dealing with small samples forecasting. The first applicable method is the Grey forecasting model. This model uses accumulative generation operator (AGO) of the Grey theory to deal with raw sample set for improving the accuracy of forecasting [10]. A second method is the use of Virtual Sample generation (VSG) [10]. This latter method fills the information gaps among each raw sample to stabilize the forecasting performance through adding newly generated virtual samples. Some examples include mega trend diffusion (MTD) [11], generalized trend diffusion (GTD) [12], genetic algorithm based virtual sample generation (GAVSG) [13] and Gaussian distribution based on VSG [14]. A particular limitation of this approach was highlighted in a study on the data generated by Generative Adversarial Networks (GANS) that showed that Top-1 and Top-5 accuracy error increased by 120% and 384% respectively and the improvement on the classification accuracy was not significant [15]. Based on this result, it was concluded that the use of such technique may incur more damage than help especially in scenarios where

 ISBN: 978-1-61208-916-4

life might be on the line. A third approach for dealing with small sample size focuses on feature extraction. This involves dimension reduction whereby useful attributes or features are identified and selected to enhance the accuracy of the forecast. During the initial period of disease outbreaks, all information acquired should be used because information on the new disease is limited. The different factors which affect the spread of the disease might not be known and as such feature extraction does not seem to be a suitable approach [10].

Based on these findings, the aim of this paper is to investigate the applicability of the Grey forecasting algorithm (GM(1,1)) to the first take off period of the coronavirus outbreak. The rest of this paper is organized as follows. Section II describes shortly the state of the art in approaches to forecasting applied to epidemics. Section III outlines the sources for the data we used for out research along with the reason for such a choice and a short description of the evaluation criteria for the assessment of the performance of forecasting algorithms. Section IV describes a few common approaches to forecasting. Section V shows the result of our simulations, both in graphical and tabular forms, with different model and data from three different countries. Eventually, the conclusions close the article.

II. State of the Art

Most approaches to forecasting require the use of statistical or machine learning models which require large sample size for higher levels of accuracy [6]. This fact in conjunction with the fact that in the early stages of epidemics data is limited inspired the application of the Grey system theory to early-stage epidemic forecasting. A great deal of research has been successfully conducted to verify the application of the Grey system theory in case of epidemics. However, to our knowledge, very little research has been carried out to compare the performance of the Grey models to other existing models like PNNs or Growth models. In this paper, we aim to shed more light on the Grey forecasting model in hope to better equip us to fight against future epidemics.

III. Methodology

In this section, we outline the sources for the data we used for out research along with the reason for such a choice and a short description of the evaluation criteria for the assessment of the performance of forecasting algorithms.

A. Data Source

Due to the novelty and the united approach society undertook against the novel coronavirus (Covid-19) pandemic, there was an ample amount of data sources. In our study, we rely on secondary data sources resulting in time saving from data collection and aggregation. We considered various options that were referred to by reliable sources such as the WHO, the European Centre for Disease, and the John Hopkins University. We decided to not use WHO data as they had shifted the timing on their reports causing overlaps in their data timelines. The data from the European Centre for Disease (ECDC) and John Hopkins data was fairly similar instead. Eventually we chose the data available on the ECDC because it included a few pre-calculated variables which made the dataset more complete for out study. The ECDC data is available on ourworldindata.org. The website and the data were compiled with the mindset that approaching a common problem together is the best way to solve it. The website confirms the daily numbers of Covid-19 cases and provides various visualizations [16].

B. Sample

A sample size of 10 was chosen to reflect more realistic situations. The WHO announced the virus's potential after eleven days compared with only 6 – 8 data points used in other studies [4][17]. We chose the data from three specific countries after considering various factors like political influences, infrastructure capabilities, corruption, and coronavirus timings. Upon thorough investigation countries which were infected early by coronavirus appear to have been caught unprepared and so the data observed from such countries appear inconsistent and un-reliable. The three countries whose data we chose to use in our study are Albania, Haiti, and Argentina. These countries were severely hit by the virus after March 2021, that means at a time when most unaffected countries were expecting and preparing for the arrival of Covid-19. Ultimately, this resulted in ensuring an adequate reporting of cases.

C. Evaluation Criteria

To assess the performance of the forecasting algorithms, there are various criterions applicable such as the Mean Average Error (MAE), the Root Mean Square Error (RMSE), and the Mean Absolute Percentage Error (MAPE). We chose to use the MAPE because of its more intuitive interpretation.

D. Systematic Review

We carried out a systematic review employing the PRISMA guidelines [18] to further understand the different components of the Grey forecasting algorithm, how to optimize these components to increase the forecasting accuracy of the Grey forecasting model.

IV. Models

In this section, we provide an overview of a few common approaches to forecasting, including the Grey system model theory, Polynomial Neural Networks, and the Gompertz model.

A. Grey System Theory

The Grey system theory works on systems with partially identified and partially unknown information by drawing out crucial information through the generation and development of the partially known information. It can describe correctly and effectively the systemic operational behavior of many

 ISBN: 978-1-61208-916-4

systems such as social, economic, agricultural, industrial, ecological, and biological systems [19]. The Grey system theory uncovers laws of change by mining and structuring available raw data, thereby representing an approach of finding data out of data. This approach is referred to as Grey sequence generation. This method considers that even though the expression of an independent system might be complex, its data is chaotic. There must be some internal laws ruling the existence of the system and its operation. Hidden laws are found through the generation of Grey numbers or functions of sequence operator. These operators allow the Grey system to handle various attributes that may be present within the data to adequately extract the information embedded within the sequence. Some of these operators include buffer operators (e.g., weakening operator, strengthening operator), average generation operators, stepwise ratio generation operators, inverse accumulating generation operators and accumulating generation operators [20].

In most instances where the GM(1,1) algorithm has been applied in research, it has almost always been an enhanced version. The most common of enhanced parameters are background values and accumulative generative operators. Background values theoretically have two primary functions. The first function is to smooth the data and strip it of its randomness and the second is to emphasize the importance of the newest datum [19]. The precision and the prediction power of the Grey model depends on the accuracy of a few coefficients (called the a and b coefficients), whereby the values of these parameters depend on the original data sequence and the background value, again reflecting the importance of the background value. Hence by improving the background value, the a and b coefficients are also improved having a compounding effect on the accuracy and performance of the model. The traditional GM(1,1) models background value is always a default value of 0.5 which is a high simplification of the modelling process meaning no preference is given to old or new data. More realistically, the value of the background value coefficient should be calculated based on the original dataset. The Accumulative Generative Operational algorithm (AGO) and the Inverse Cumulative Generative Operational algorithm (I-AGO) are important parts of the Grey model algorithm and are responsible for reducing the variation or fluctuations of the original series. The problem found in the algorithms used for the AGO sequence generation comes from the definition of class ratios when creating the Grey model algorithm by assuming that each component of the original sequence is positive and the class ratio of the first AGO manipulation is always smaller than the class ratio of the second AGO manipulation [20]. This means that in some cases the first accumulated generating operator violates the principle of new information priority and principle of minimal information of Grey system theory. However, it can be proved that the condition stated above is only sufficient but is not a necessary one [21].

B. *Polynomial Neural Network*

A previous study investigated the early-stage epidemic forecasting and concluded that the best algorithm is the Polynomial neural networks with corrected feedback with an RMSE of 136 [4]. PNNs belong to a group of neural algorithms whose theory roots back to and is based on works from the early 1970s [22]. Those types of neural networks are "self-organizing" networks to express the fact that the connections between the neurons in the network are selected during the training phase to optimize the network. The number of layers in the network are also determined automatically to produce maximum accuracy without overfitting [23].

C. *Gompertz Model*

Growth models are currently the most used methods for early-stage forecasting of disease outbreaks. Recently, these models were applied to twenty infectious disease outbreaks representing a range of transmission routes and a wide range of epidemic profiles. The outbreaks ranged from slow growth (sub-exponential) to fast growth (close to exponential). The proposed generalized growth model outperformed the other algorithms that it was compared to [6].

Growth models can be grouped into at least two main categories. The first group includes models without inflection points such as the Brody and negative models [24]. The second group include models with sigmoidal shape that also poses a fixed inflection points such as the Gompertz, the logistic or the von Bertalanffy models [24]. The Gompertz and the logistic models are the most frequently adopted models in literature. Given that the cumulative number of cases by coronavirus appears to form an asymmetrical sigmoidal curve, the logistic and the Gompertz model naturally lend themselves as ideal candidates to apply [24].

The Gompertz model and logistic model are both growth models with similar properties, making them useful for the representation of the generalized growth models. There has not been a distinct advantage found between both models as both require constants corresponding to upper asymptote, time origin, and time unit or rate [25]. But in practice, it has been found that the logistic model gives good fits on material showing an inflection about midway between the asymptotes, while the inflection point for the Gompertz model is about 37%. This makes the Gompertz model more suitable for first take off period forecasting [25].

D. *Average Ensemble*

Enhancing the traditional GM(1,1) algorithm can be achieved in multiple ways, all of which offer different advantages and consequently have different effects on the algorithm. Based on the systematic review we conducted, the use of multi-model approach (ensembles) turned out to be one of the most popular methods and proved to be the most suitable method to apply in this paper. There are various types of ensembles methods namely bagging, boosting, Adaboost, random forest, gradient boosting, averages etc. [26].

For the purpose of this paper, we applied the average ensemble to the Covid-19 dataset in order to demonstrate the application of one of the methods of enhancement. Using the average of forecasts is one of the simple but effective methods of ensembles. Previous studies [27] suggest that using averages of forecasts provides improved forecasting accuracy and that the variability of accuracy among different combinations decreases, as the number of methods in the average increases. Similar earlier studies also indicate that, in some cases, the simple average ensemble can outperform its weighted average counterpart [28].

In the next section, we show the forecasting results we obtained using the coronavirus dataset of Albania, Argentina, and Haiti. The biggest influence on the countries chosen is the quality of data generated. Upon thorough observations of the data, countries like China who were hit early by the virus seem to have been un-prepared meaning the necessary actions required to adequately record and publish the coronavirus numbers were not in place and so countries which contracted coronavirus after March 2020 were chosen as their published data appear more reliable and consistent.

V. RESULTS

In this section we show the result of our simulations with different model for a few selected countries.

A. Albania Forecast

Figure 1 shows how well each model describes the Albania dataset and forecasts the Covid-19 cases. The y axis shows the number of cases while the x axis shows the days that have elapsed. The red line shows the actual data. The other lines show the fit and forecast for the models we chose to employ, namely the GM(1,1), the Gompertz, the PNN, and the ensemble model. The graphics indicates that the best models for the Albania dataset are the Gompertz model and the PNN while the GM(1,1) appears to be the worst fitting one. Table I shows the MAPE values for all models we considered in out simulations when the data for Argentina was used.

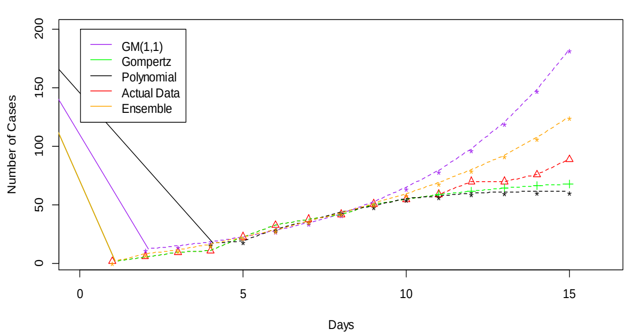

Figure 1. Model values for GM(1,1), Gompertz, PNN, Ensemble model for the Albania Dataset.

TABLE I. MAPE VALUES FOR ALL MODELS AND THE DATA SET FOR ALBANIA.

	In Sample	*Out of Sample*
GM(1,1)	30.67	69.76
Gompertz	10.50	11.16
PNN	13.16	15.53
Ensemble	15.92	29.35

B. Argentina Forecast

Figure 2 visually highlights the performance of the four different models applied to the dataset for the country Argentina. The y axis shows the number of cases while the x axis shows the days that have elapsed. Again, the red line shows the actual data available for Argentina. The GM(1,1) seems to forecast very high values, which is clear when observing the purple and the red line. The best performing model in this case appears indistinguishable but this can be explained by comparing the MAPE values as seen in Table II.

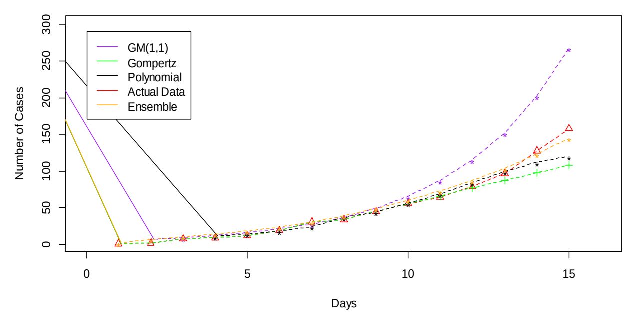

Figure 2. Forecast values for GM(1,1), Gompertz, PNN, Ensemble model for the Argentina Dataset.

TABLE II. MAPE VALUES FOR ALL MODELS AND THE DATA SET FOR ARGENTINA.

	In Sample	*Out of Sample*
GM(1,1)	43.28	53.11
Gompertz	18.69	13.90
PNN	12.24	10.27
Ensemble	54.67	8.47

C. Haiti Forecast

Figure 3 highlights visually how well the models fit and forecasts the Haiti dataset. In this case, the Gompertz model appears to have a particularly good fit for the data set, but aside from this, it can be difficult to distinguish and discriminate using this graph only. Table III outlines the MAPE values and thus can provide a clearer understanding on model accuracy for the data related to Haiti.

 ISBN: 978-1-61208-916-4

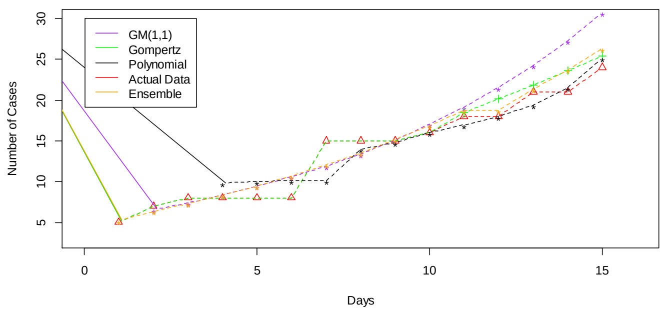

Figure 3. Forecast values for GM(1,1), Gompertz, PNN, Ensemble model for the Haiti Dataset.

TABLE III. MAPE VALUES FOR ALL MODELS AND THE DATA SET OF HAITI.

	In Sample	*Out of Sample*
GM(1,1)	11.73	19.90
Gompertz	11.83	7.33
PNN	16.67	4.15
Ensemble	11.43	6.53

As already mentioned, the MAPE values give a better intuitive understanding of forecast accuracy. Overall, the Gompertz models and the PNN both perform equally well on different datasets suggesting the need for further tests using other datasets. The GM(1,1) consistently performed as the worst. While the ensemble was implemented to show the potential for improving the GM(1,1), the improvement can be attributed to the outstanding performance of the Gompertz model.

D. Systematic Review

These results from the experiments initiated an investigation into the methods which the Grey forecasting model's parameters could be optimized to provide better accuracy. From the studies that we reviewed systematically, 38% improved the background value of the Grey forecasting model, 23% used a multi-model approach similar to an ensemble while 15% enhanced the Grey forecasting model GM(1,1) by improving the accumulative generating operation mechanism, other methods employed including the use of Virtual sample generation (VSG), parameter optimization, and finally the diffusion models.

VI. Conclusion

The comparison of the GM(1,1) forecasting algorithm to the PNN and Gompertz model against the datasets showed that the traditional GM(1,1) algorithm was not able to match the forecasting capabilities of the Gompertz and PNN model. This result was unexpected but has been attributed to the sub-exponential nature of the dataset and the hidden nonoptimized parameters within the GM(1,1) algorithm. While we see that the enhanced ensemble GM(1,1) algorithm performs better in comparison to the traditional GM(1,1), enhanced GM(1,1) models which optimize the parameters within the GM(1,1) model itself have immense potential in the forecasting of early stage epidemic forecasting.

The preliminary conclusions of this paper are confirmed in a separate research study by Liu and colleagues [29]. Liu uses fractional Grey model FGM(1,1) that is an enhanced version of the traditional forecasting model GM(1,1) by using a fractional order accumulation. Liu's FGM(1,1) model can better adhere to the principle of new information priority and was applied to forecast the first take off period of Covid-19 with good results. In another study [30], the GM(1,1) model was also applied to and compared with other modifications such as the non-linear Grey Bernoulli model (NGBM(1,1) and the FGM(1,1) model to long term cumulative forecasting of the Covid-19 epidemic cases in UK, USA and Italy. The values of root mean square error, absolute percentage error (APE) and R^2 values for GM (1,1), NGBM (1,1) and fractional accumulative nonlinear Grey Bernoulli model (FANGBM(1,1)) for the prediction of the cumulative cases of Covid-19 indicate that the FANGBM is the most accurate with the highest R^2 and lowest RMSE and APE values.

Based on all this evidence, the aim of this paper to confirm the applicability of Grey forecasting model to forecast early-stage epidemic cases has been achieved.

References

[1] World Health Organisation (WHO) (2020) Coronavirus disease Weekly Epidemiological Update and Weekly Operational Update https://www.who.int/emergencies/diseases/novel-coronavirus-2019/ [Accessed 02/09/2020].

[2] Unkel, F., Farrington, P., & Garthwaite, P., (2012) Statistical Methods for the Prospective Detection of Infectious Disease Outbreaks: a review. *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, Vol.175, Nr. 1, pp. 49-82.

[3] Tabataba, F.S, Prithwish, C., Ramakrishman, N., Venkatramanan, S., Jiangzhou, C., Lewis, B., & Marathe, M. (2017) A Framework for evaluating epidemic forecasts, *BMC Infectious Diseases* Vol. 17, Nr. 345.

[4] Fong, S.J., Li, G., Dey, N., Crepo, R.G., & Herrera-Viedma, E. (2020) Finding an Accurate Early Forecasting model from small dataset: A case of 2019-ncov novel coronavirus outbreak. *International Journal of Interactive Multimedia and Artificial Intelligence*, Vol. 6. Nr. 1, pp. 132-40

[5] Viboud, C., Simonsen, L., & Chowell, G., (2016) A generalised growth model to characterise the early ascending phase of infectious disease. *Epidemics*, Vol. 15, June 2016, pp 27-37.

[6] Hyndman, R.J., & Athanasopoulos, G. (2018) *Forecasting: principles and practice*, 2nd edition, OTexts: Melbourne, Australia. OTexts.com/fpp2. [Accessed 02/09/2020]

[7] Zhu, B., Chen, Z., & Le, Y., (2016). A novel mega-trend-diffusion for small sample. *CIESC Journal*, Vol. 67, pp. 820-826.

[8] Chang, C.J., Li, D.C., Dai, W.L., & Chen, C.C. (2014) A Latent function to extend domain attributes to improve the accuracy of small data set forecasting. *Neurocomputing*, Vol. 129, April 2014, pp. 343–349.

[9] Li, D.C., Fang, Y.H., Lai, Y.Y., & Hu, S.C. (2009) Utilization of virtual sample to facilitate cancer identification for DNA microarray data in the early stages of an investigation. *Information Sciences*, Vol. 179, Nr. 16, pp. 2740-2753.

[10] Chen, Z., Zhu, B., He, Y., & Yu, L. (2017) A PSO based virtual sample generation method for small sample sets: Applications to regression datasets. *Engineering Applications of Artificial Intelligence*, Vol. 59, March 2017, Pages 236-243.

[11] Li, D.C., Wu, C.S., Sai, T.I., & Lina, Y.S. (2007) Using mega-trend diffusion and artificial samples in small dataset for early flexible manufacturing system scheduling knowledge. *Computers & Operations Research*, Vol. 34, Nr. 4, pp. 966-982.

[12] Lin, Y.S., & Li, D.C. (2010) The generalization trend-diffusion modeling algorithm for small datasets in the early stages of manufacturing systems. *European Journal of Operational Research* Vol. 207, Nr. 1, pp 121-130.

[13] Li, D.C., & Wen, I.H. (2014) A genetic based virtual sample generation technique to improve small data set learning. *Neurocomputing*, Vol. 143, Nr., pp. 222-230.

[14] Yang, J., Yu, X., Xie, Z., & Zhang, J. (2011) A novel virtual sample generation method based on Gaussian distribution. *Knowledge-Based Systems*, Vol. 24, Nr. 6, pp. 740-748.

[15] Ravuri, S., & Vinyals, O. (2019) Seeing is not necessarily Believing : Limitations of BigGans for Data Augmentation.

[16] Ritchie, H., Mathieu, E., Rodés-Guirao, L., Appel, C., Giattino, C. Ortiz-Ospina, E., Hasell, J., Macdonald, B., Beltekian, D., & Roser, M. (2020) Coronavirus Pandemic (COVID-19). Published online at: https://ourworldindata.org/coronavirus [Accessed 02/09/2020]

[17] Wu, K., Darcet, D., Wang, Q., & Sornette, D. (2020) Generalised Logistic growth modeling of the COVID-19 outbreak in 29 provinces in China and in the rest of the world. *Nonlinear Dynamics*, Vol. 101, pp. 1561–1581.

[18] Page, M.J., McKenzie, J.E., Bossuyt, P.M., Boutron, I., Hoffmann, T.C., & Mulrow, C.D. (2020). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*, March 2021, Nr. 372.

[19] Lin, Y., & Liu, S. (2004) A Historical Introduction to Grey systems Theory. In: Proceedings of IEEE International Conference on Systems, Man and Cybernetics, Vol.3, pp. 2403-2408.

[20] Deng, J.L. (1982) Control problems of Grey systems, *Systems & Control Letters*, Vol. 1 No. 5, pp. 288-94.

[21] Wen, K.-L., Dai, J.-W., Chang, T.-C., Tong, C-C. (2001) The discussions of class ratio for AGO algorithm in Grey theory. In: Proceedings of IEEE International Conference on Systems, Man and Cybernetics, Vol. 1, pp. 7-11.

[22] Ivakhnenko, A.G. (1971) Ploynomial Theory of Complex Systems. IEEE Transactions on Systams, Man, and cybernetics, Vol. 1, Nr. 4, pp. 364-378.

[23] Dag, O., Yozgatligil, C. (2016) GMDH: An R Package for Short-Term Forecasting via GMDH-Type Neural Network Algorithms. *The R Journal*, Vol. 8, Nr. 1, pp. 379-386.

[24] Mandujano Valle, J.A. (2020) Predicting the number of total COVID-19 cases and deaths in Brazil by the Gompertz model. *Nonlinear Dynamics*, Vol. 102, pp. 2951–2957.

[25] Winsor, C.O. (1931) The Gompertz Curve as a Growth Curve. *Proceedings of the Natlional Academy of Science USA*, Vol. 18, Nr. 1, pp. 1–8.

[26] Latka, M., Ignaccolo, M., Jernajczyk, W., & West, B. (2010). Time and ensemble averaging in time series analysis.

[27] Makridakis, S., & Winkler, R. L. (1983) Averages of Forecasts - Some Empirical Results, *Management Science*, Vol. 29, pp. 987-996.

[28] Palm, F. C. , & Zellner, A. (1992) To combine or not to combine - Issues of combining forecasts, *Journal of Forecasting*, vol. 11, pp. 687-701.

[29] Liu, L., Chen, Y,. & Wu, L. (2020) Forecasting Confirmed Cases, Deaths, and Recoveries from COVID-19 in China during the Early Stage. Mathematical Problems in Engineering. 1-4. 10.

[30] Şahin, U. & Şahin, T. (2020) Forecasting the cumulative number of confirmed cases of COVID-19 in Italy, UK and USA using fractional nonlinear Grey Bernoulli model. *Chaos Solitons & Fractals*, Vol. 138.